Polynucleotide and deduced amino acid sequence of hMLH1

```
-40                 -20                  1
.         .         .         .         .         .
GTTGAACATCTAGACGTTTCCTTGGCTCTTCTGGCGCCAAAATGTCGTTCGTGGCAGGGG
+---------+---------+---------+---------+---------+---------
CAACTTGTAGATCTGCAAAGGAACCGAGAAGACCGCGGTTTTACAGCAAGCACCGTCCCC
                                        M  S  F  V  A  G  V
20                  40                  60
.         .         .         .         .         .
TTATTCGGCGGCTGGACGAGACAGTGGTGAACCGCATCGCGGCGGGGGAAGTTATCCAGC
+---------+---------+---------+---------+---------+---------
AATAAGCCGCCGACCTGCTCTGTCACCACTTGGCGTAGCGCCGCCCCCTTCAATAGGTCG
  I  R  R  L  D  E  T  V  V  N  R  I  A  A  G  E  V  I  Q  R
80                  100                 120
.         .         .         .         .         .
GGCCAGCTAATGCTATCAAAGAGATGATTGAGAACTGTTTAGATGCAAAATCCACAAGTA
+---------+---------+---------+---------+---------+---------
CCGGTCGATTACGATAGTTTCTCTACTAACTCTTGACAAATCTACGTTTTAGGTGTTCAT
  P  A  N  A  I  K  E  M  I  E  N  C  L  D  A  K  S  T  S  I
140                 160                 180
.         .         .         .         .         .
TTCAAGTGATTGTTAAAGAGGGAGGCCTGAAGTTGATTCAGATCCAAGACAATGGCACCG
+---------+---------+---------+---------+---------+---------
AAGTTCACTAACAATTTCTCCCTCCGGACTTCAACTAAGTCTAGGTTCTGTTACCGTGGC
  Q  V  I  V  K  E  G  G  L  K  L  I  Q  I  Q  D  N  G  T  G
200                 220                 240
.         .         .         .         .         .
GGATCAGGAAAGAAGATCTGGATATTGTATGTGAAAGGTTCACTACTAGTAAACTGCAGT
+---------+---------+---------+---------+---------+---------
CCTAGTCCTTTCTTCTAGACCTATAACATACACTTTCCAAGTGATGATCATTTGACGTCA
  I  R  K  E  D  L  D  I  V  C  E  R  F  T  T  S  K  L  Q  S
260                 280                 300
.         .         .         .         .         .
CCTTTGAGGATTTAGCCAGTATTTCTACCTATGGCTTTCGAGGTGAGGCTTTGGCCAGCA
+---------+---------+---------+---------+---------+---------
GGAAACTCCTAAATCGGTCATAAAGATGGATACCGAAAGCTCCACTCCGAAACCGGTCGT
  F  E  D  L  A  S  I  S  T  Y  G  F  R  G  E  A  L  A  S  I
320                 340                 360
.         .         .         .         .         .
TAAGCCATGTGGCTCATGTTACTATTACAACGAAAACAGCTGATGGAAAGTGTGCATACA
+---------+---------+---------+---------+---------+---------
ATTCGGTACACCGAGTACAATGATAATGTTGCTTTTGTCGACTACCTTTCACACGTATGT
  S  H  V  A  H  V  T  I  T  T  K  T  A  D  G  K  C  A  Y  R
```

FIG. 1A

```
380                 400                 420
.         .         .         .         .         .
GAGCAAGTTACTCAGATGGAAAACTGAAAGCCCCTCCTAAACCATGTGCTGGCAATCAAG
+---------+---------+---------+---------+---------+---------+---------
CTCGTTCAATGAGTCTACCTTTTGACTTTCGGGGAGGATTTGGTACACGACCGTTAGTTC
  A  S  Y  S  D  G  K  L  K  A  P  P  K  P  C  A  G  N  Q  G
440                 460                 480
.         .         .         .         .         .
GGACCCAGATCACGGTGGAGGACCTTTTTTACAACATAGCCACGAGGAGAAAAGCTTTAA
+---------+---------+---------+---------+---------+---------+---------
CCTGGGTCTAGTGCCACCTCCTGGAAAAAATGTTGTATCGGTGCTCCTCTTTTCGAAATT
  T  Q  I  T  V  E  D  L  F  Y  N  I  A  T  R  R  K  A  L  K
500                 520                 540
.         .         .         .         .         .
AAAATCCAAGTGAAGAATATGGGAAAATTTTGGAAGTTGTTGGCAGGTATTCAGTACACA
+---------+---------+---------+---------+---------+---------+---------
TTTTAGGTTCACTTCTTATACCCTTTTAAAACCTTCAACAACCGTCCATAAGTCATGTGT
  N  P  S  E  E  Y  G  K  I  L  E  V  V  G  R  Y  S  V  H  N
560                 580                 600
.         .         .         .         .         .
ATGCAGGCATTAGTTTCTCAGTTAAAAAACAAGGAGAGACAGTAGCTGATGTTAGGACAC
+---------+---------+---------+---------+---------+---------+---------
TACGTCCGTAATCAAAGAGTCAATTTTTTGTTCCTCTCTGTCATCGACTACAATCCTGTG
  A  G  I  S  F  S  V  K  K  Q  G  E  T  V  A  D  V  R  T  L
620                 640                 660
.         .         .         .         .         .
TACCCAATGCCTCAACCGTGGACAATATTCGCTCCGTCTTTGGAAATGCTGTTAGTCGAG
+---------+---------+---------+---------+---------+---------+---------
ATGGGTTACGGAGTTGGCACCTGTTATAAGCGAGGCAGAAACCTTTACGACAATCAGCTC
  P  N  A  S  T  V  D  N  I  R  S  V  F  G  N  A  V  S  R  E
680                 700                 720
.         .         .         .         .         .
AACTGATAGAAATTGGATGTGAGGATAAAACCCTAGCCTTCAAAATGAATGGTTACATAT
+---------+---------+---------+---------+---------+---------+---------
TTGACTATCTTTAACCTACACTCCTATTTTGGGATCGGAAGTTTTACTTACCAATGTATA
  L  I  E  I  G  C  E  D  K  T  L  A  F  K  M  N  G  Y  I  S
740                 760                 780
.         .         .         .         .         .
CCAATGCAAACTACTCAGTGAAGAAGTGCATCTTCTTACTCTTCATCAACCATCGTCTGG
+---------+---------+---------+---------+---------+---------+---------
GGTTACGTTTGATGAGTCACTTCTTCACGTAGAAGAATGAGAAGTAGTTGGTAGCAGACC
  N  A  N  Y  S  V  K  K  C  I  F  L  L  F  I  N  H  R  L  V
```

FIG. 1B

```
800                 820                 840
.        .         .         .         .         .
TAGAATCAACTTCCTTGAGAAAAGCCATAGAAACAGTGTATGCAGCCTATTTGCCCAAAA
+---------+---------+---------+---------+---------+---------
ATCTTAGTTGAAGGAACTCTTTTCGGTATCTTTGTCACATACGTCGGATAAACGGGTTTT
  E  S  T  S  L  R  K  A  I  E  T  V  Y  A  A  Y  L  P  K  N
860                 880                 900
.        .         .         .         .         .
ACACACACCCATTCCTGTACCTCAGTTTAGAAATCAGTCCCCAGAATGTGGATGTTAATG
+---------+---------+---------+---------+---------+---------
TGTGTGTGGGTAAGGACATGGAGTCAAATCTTTAGTCAGGGGTCTTACACCTACAATTAC
  T  H  P  F  L  Y  L  S  L  E  I  S  P  Q  N  V  D  V  N  V
920                 940                 960
.        .         .         .         .         .
TGCACCCCACAAAGCATGAAGTTCACTTCCTGCACGAGGAGAGCATCCTGGAGCGGGTGC
+---------+---------+---------+---------+---------+---------
ACGTGGGGTGTTTCGTACTTCAAGTGAAGGACGTGCTCCTCTCGTAGGACCTCGCCCACG
  H  P  T  K  H  E  V  H  F  L  H  E  E  S  I  L  E  R  V  Q
980                 1000                1020
.        .         .         .         .         .
AGCAGCACATCGAGAGCAAGCTCCTGGGCTCCAATTCCTCCAGGATGTACTTCACCCAGA
+---------+---------+---------+---------+---------+---------
TCGTCGTGTAGCTCTCGTTCGAGGACCCGAGGTTAAGGAGGTCCTACATGAAGTGGGTCT
  Q  H  I  E  S  K  L  L  G  S  N  S  S  R  M  Y  F  T  Q  T
1040                1060                1080
.        .         .         .         .         .
CTTTGCTACCAGGACTTGCTGGCCCCTCTGGGGAGATGGTTAAATCCACAACAAGTCTGA
+---------+---------+---------+---------+---------+---------
GAAACGATGGTCCTGAACGACCGGGGAGACCCCTCTACCAATTTAGGTGTTGTTCAGACT
  L  L  P  G  L  A  G  P  S  G  E  M  V  K  S  T  T  S  L  T
1100                1120                1140
.        .         .         .         .         .
CCTCGTCTTCTACTTCTGGAAGTAGTGATAAGGTCTATGCCCACCAGATGGTTCGTACAG
+---------+---------+---------+---------+---------+---------
GGAGCAGAAGATGAAGACCTTCATCACTATTCCAGATACGGGTGGTCTACCAAGCATGTC
  S  S  S  T  S  G  S  S  D  K  V  Y  A  H  Q  M  V  R  T  D
1160                1180                1200
.        .         .         .         .         .
ATTCCCGGGAACAGAAGCTTGATGCATTTCTGCAGCCTCTGAGCAAACCCCTGTCCAGTC
+---------+---------+---------+---------+---------+---------
TAAGGGCCCTTGTCTTCGAACTACGTAAAGACGTCGGAGACTCGTTTGGGGACAGGTCAG
  S  R  E  Q  K  L  D  A  F  L  Q  P  L  S  K  P  L  S  S  Q
```

FIG. 1C

```
1220                1240                1260
.         .         .         .         .         .
AGCCCCAGGCCATTGTCACAGAGGATAAGACAGATATTTCTAGTGGCAGGGCTAGGCAGC
+---------+---------+---------+---------+---------+---------
TCGGGGTCCGGTAACAGTGTCTCCTATTCTGTCTATAAAGATCACCGTCCCGATCCGTCG
  P  Q  A  I  V  T  E  D  K  T  D  I  S  S  G  R  A  R  Q  Q
1280                1300                1320
.         .         .         .         .         .
AAGATGAGGAGATGCTTGAACTCCCAGCCCCTGCTGAAGTGGCTGCCAAAAATCAGAGCT
+---------+---------+---------+---------+---------+---------
TTCTACTCCTCTACGAACTTGAGGGTCGGGGACGACTTCACCGACGGTTTTTAGTCTCGA
  D  E  E  M  L  E  L  P  A  P  A  E  V  A  A  K  N  Q  S  L
1340                1360                1380
.         .         .         .         .         .
TGGAGGGGGATACAACAAAGGGGACTTCAGAAATGTCAGAGAAGAGAGGACCTACTTCCA
+---------+---------+---------+---------+---------+---------
ACCTCCCCCTATGTTGTTTCCCCTGAAGTCTTTACAGTCTCTTCTCTCCTGGATGAAGGT
  E  G  D  T  T  K  G  T  S  E  M  S  E  K  R  G  P  T  S  S
1400                1420                1440
.         .         .         .         .         .
GCAACCCCAGAAAGAGACATCGGGAAGATTCTGATGTGGAAATGGTGGAAGATGATTCCC
+---------+---------+---------+---------+---------+---------
CGTTGGGGTCTTTCTCTGTAGCCCTTCTAAGACTACACCTTTACCACCTTCTACTAAGGG
  N  P  R  K  R  H  R  E  D  S  D  V  E  M  V  E  D  D  S  R
1460                1480                1500
.         .         .         .         .         .
GAAAGGAAATGACTGCAGCTTGTACCCCCCGGAGAAGGATCATTAACCTCACTAGTGTTT
+---------+---------+---------+---------+---------+---------
CTTTCCTTTACTGACGTCGAACATGGGGGGCCTCTTCCTAGTAATTGGAGTGATCACAAA
  K  E  M  T  A  A  C  T  P  R  R  R  I  I  N  L  T  S  V  L
1520                1540                1560
.         .         .         .         .         .
TGAGTCTCCAGGAAGAAATTAATGAGCAGGGACATGAGGTTCTCCGGGAGATGTTGCATA
+---------+---------+---------+---------+---------+---------
ACTCAGAGGTCCTTCTTTAATTACTCGTCCCTGTACTCCAAGAGGCCCTCTACAACGTAT
  S  L  Q  E  E  I  N  E  Q  G  H  E  V  L  R  E  M  L  H  N
1580                1600                1620
.         .         .         .         .         .
ACCACTCCTTCGTGGGCTGTGTGAATCCTCAGTGGGCCTTGGCACAGCATCAAACCAAGT
+---------+---------+---------+---------+---------+---------
TGGTGAGGAAGCACCCGACACACTTAGGAGTCACCCGGAACCGTGTCGTAGTTTGGTTCA
  H  S  F  V  G  C  V  N  P  Q  W  A  L  A  Q  H  Q  T  K  L
```

FIG. 1D

```
1640                1660                1680
.         .         .         .         .         .
TATACCTTCTCAACACCACCAAGCTTAGTGAAGAACTGTTCTACCAGATACTCATTTATG
+---------+---------+---------+---------+---------+---------+
ATATGGAAGAGTTGTGGTGGTTCGAATCACTTCTTGACAAGATGGTCTATGAGTAAATAC
  Y  L  L  N  T  T  K  L  S  E  E  L  F  Y  Q  I  L  I  Y  D
1700                1720                1740
.         .         .         .         .         .
ATTTTGCCAATTTTGGTGTTCTCAGGTTATCGGAGCCAGCACCGCTCTTTGACCTTGCCA
+---------+---------+---------+---------+---------+---------+
TAAAACGGTTAAAACCACAAGAGTCCAATAGCCTCGGTCGTGGCGAGAAACTGGAACGGT
  F  A  N  F  G  V  L  R  L  S  E  P  A  P  L  F  D  L  A  M
1760                1780                1800
.         .         .         .         .         .
TGCTTGCCTTAGATAGTCCAGAGAGTGGCTGGACAGAGGAAGATGGTCCCAAAGAAGGAC
+---------+---------+---------+---------+---------+---------+
ACGAACGGAATCTATCAGGTCTCTCACCGACCTGTCTCCTTCTACCAGGGTTTCTTCCTG
  L  A  L  D  S  P  E  S  G  W  T  E  E  D  G  P  K  E  G  L
1820                1840                1860
.         .         .         .         .         .
TTGCTGAATACATTGTTGAGTTTCTGAAGAAGAAGGCTGAGATGCTTGCAGACTATTTCT
+---------+---------+---------+---------+---------+---------+
AACGACTTATGTAACAACTCAAAGACTTCTTCTTCCGACTCTACGAACGTCTGATAAAGA
  A  E  Y  I  V  E  F  L  K  K  K  A  E  M  L  A  D  Y  F  S
1880                1900                1920
.         .         .         .         .         .
CTTTGGAAATTGATGAGGAAGGGAACCTGATTGGATTACCCCTTCTGATTGACAACTATG
+---------+---------+---------+---------+---------+---------+
GAAACCTTTAACTACTCCTTCCCTTGGACTAACCTAATGGGGAAGACTAACTGTTGATAC
  L  E  I  D  E  E  G  N  L  I  G  L  P  L  L  I  D  N  Y  V
1940                1960                1980
.         .         .         .         .         .
TGCCCCCTTTGGAGGGACTGCCTATCTTCATTCTTCGACTAGCCACTGAGGTGAATTGGG
---------+---------+---------+---------+---------+---------+
ACGGGGGAAACCTCCCTGACGGATAGAAGTAAGAAGCTGATCGGTGACTCCACTTAACCC
  P  P  L  E  G  L  P  I  F  I  L  R  L  A  T  E  V  N  W  D
2000                2020                2040
.         .         .         .         .         .
ACGAAGAAAAGGAATGTTTTGAAAGCCTCAGTAAAGAATGCGCTATGTTCTATTCCATCC
+---------+---------+---------+---------+---------+---------+
TGCTTCTTTTCCTTACAAAACTTTCGGAGTCATTTCTTACGCGATACAAGATAAGGTAGG
  E  E  K  E  C  F  E  S  L  S  K  E  C  A  M  F  Y  S  I  R
```

FIG. 1E

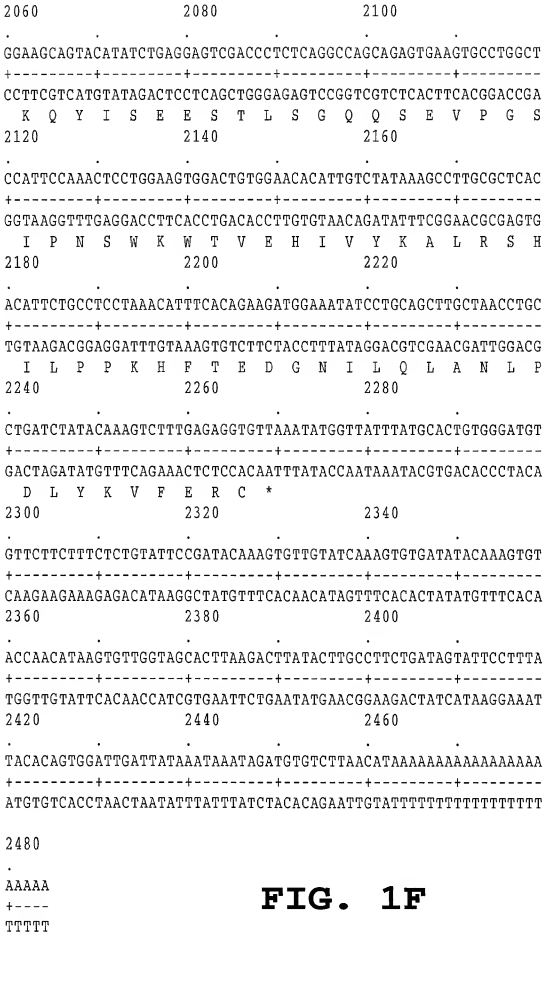
2060
2080
2100
GGAAGCAGTACATATCTGAGGAGTCGACCCTCTCAGGCCAGCAGAGTGAAGTGCCTGGCT
CCTTCGTCATGTATAGACTCCTCAGCTGGGAGAGTCCGGTCGTCTCACTTCACGGACCGA
K Q Y I S E E S T L S G Q Q S E V P G S
2120
2140
2160
CCATTCCAAACTCCTGGAAGTGGACTGTGGAACACATTGTCTATAAAGCCTTGCGCTCAC
GGTAAGGTTTGAGGACCTTCACCTGACACCTTGTGTAACAGATATTTCGGAACGCGAGTG
I P N S W K W T V E H I V Y K A L R S H
2180
2200
2220
ACATTCTGCCTCCTAAACATTTCACAGAAGATGGAAATATCCTGCAGCTTGCTAACCTGC
TGTAAGACGGAGGATTTGTAAAGTGTCTTCTACCTTTATAGGACGTCGAACGATTGGACG
I L P P K H F T E D G N I L Q L A N L P
2240
2260
2280
CTGATCTATACAAAGTCTTTGAGAGGTGTTAAATATGGTTATTTATGCACTGTGGGATGT
GACTAGATATGTTTCAGAAACTCTCCACAATTTATACCAATAAATACGTGACACCCTACA
D L Y K V F E R C *
2300
2320
2340
GTTCTTCTTTCTCTGTATTCCGATACAAAGTGTTGTATCAAAGTGTGATATACAAAGTGT
CAAGAAGAAAGAGACATAAGGCTATGTTTCACAACATAGTTTCACACTATATGTTTCACA
2360
2380
2400
ACCAACATAAGTGTTGGTAGCACTTAAGACTTATACTTGCCTTCTGATAGTATTCCTTTA
TGGTTGTATTCACAACCATCGTGAATTCTGAATATGAACGGAAGACTATCATAAGGAAAT
2420
2440
2460
TACACAGTGGATTGATTATAAATAAATAGATGTGTCTTAACATAAAAAAAAAAAAAAAAA
ATGTGTCACCTAACTAATATTTATTTATCTACACAGAATTGTATTTTTTTTTTTTTTTTT
2480
AAAAA
TTTTT

FIG. 1F

Polynucleotide and deduced amino acid sequence of hMLH2

```
        -70                 -50                 -30
         .         .         .         .         .         .
GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG
---------+---------+---------+---------+---------+---------+
CCGTGCTCACCGACGAACGCCGATCACCTACCATTAACGGACGGAGCGCGATCGTCGTTC
        -10                 10                  30
         .         .         .         .         .         .
CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA
---------+---------+---------+---------+---------+---------+
GACGAGACAATTTTCGCTTTTACTTTGTTAACGGACGCCGTTGTCAAGCTGAGGAAAGTT
                   M  K  Q  L  P  A  A  T  V  R  L  L  S  S
        50                  70                  90
         .         .         .         .         .         .
GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG
---------+---------+---------+---------+---------+---------+
CAAGAGTCTAGTAGTGAAGCCACCAGTCACAACATTTTCTCGAATAACTTTTGAGGAACC
  S  Q  I  I  T  S  V  V  S  V  V  K  E  L  I  E  N  S  L  D
        110                 130                 150
         .         .         .         .         .         .
ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG
---------+---------+---------+---------+---------+---------+
TACGACCACGGTGTTCGCATCTACAATTTGACCTCTTGATACCTAAACTATTTTAACTCC
  A  G  A  T  S  V  D  V  K  L  E  N  Y  G  F  D  K  I  E  V
        170                 190                 210
         .         .         .         .         .         .
TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT
---------+---------+---------+---------+---------+---------+
ACGCTCTATTGCCCCTCCCATAGTTCCGACAACTACGTGGACATTACCGTTACTTCATGA
  R  D  N  G  E  G  I  K  A  V  D  A  P  V  M  A  M  K  Y  Y
        230                 250                 270
         .         .         .         .         .         .
ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG
---------+---------+---------+---------+---------+---------+
TGTGGAGTTTTTATTTATCAGTACTTCTAGAACTTTTAAACTGTTGAATGCCAAAAGCAC
  T  S  K  I  N  S  H  E  D  L  E  N  L  T  T  Y  G  F  R  G
        290                 310                 330
         .         .         .         .         .         .
GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG
---------+---------+---------+---------+---------+---------+
CTCTTCGGAACCCCAGTTAAACAACATATCGACTCCAAAATTAATGTTGTTCTTGCCGAC
  E  A  L  G  S  I  C  C  I  A  E  V  L  I  T  T  R  T  A  A
```

FIG. 2A

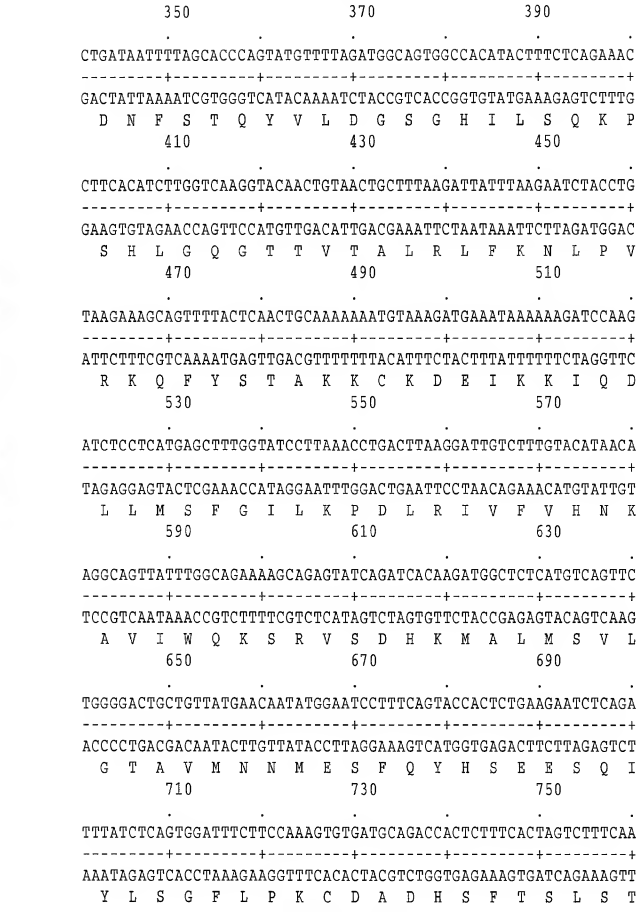

FIG. 2B

```
         770               790               810
          .         .         .         .         .         .
CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA
---------+---------+---------+---------+---------+---------+
GTGGTCTTTCTTCAAAGTAGAAGTATTTGTCAGCTGGTCATGTAGTTTTTCTATAGAATT
 P  E  R  S  F  I  F  I  N  S  R  P  V  H  Q  K  D  I  L  K
         830               850               870
          .         .         .         .         .         .
AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG
---------+---------+---------+---------+---------+---------+
TCAATTAGGCTGTAGTAATGTTAGACTTTACGGATTTCCTTAGATGAGCAAACATAGGAC
 L  I  R  H  H  Y  N  L  K  C  L  K  E  S  T  R  L  Y  P  V
         890               910               930
          .         .         .         .         .         .
TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA
---------+---------+---------+---------+---------+---------+
AAAAGAAAGACTTTTAGCTACAAGGATGTCGACTACAACTACATTTAAATTGTGGTCTAT
 F  F  L  K  I  D  V  P  T  A  D  V  D  V  N  L  T  P  D  K
         950               970               990
          .         .         .         .         .         .
AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA
---------+---------+---------+---------+---------+---------+
TTTCGGTTCATAATAATGTTTTATTCCTTAGACAAAATTAACGAGAACTTTTAGACTACT
 S  Q  V  L  L  Q  N  K  E  S  V  L  I  A  L  E  N  L  M  T
         1010              1030              1050
          .         .         .         .         .         .
CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT
---------+---------+---------+---------+---------+---------+
GCTGAACAATACCTGGTAATGGATCATGTTTAAGAATACTTTTATTATTTTGTCTACAAA
 T  C  Y  G  P  L  P  S  T  N  S  Y  E  N  N  K  T  D  V  S
         1070              1090              1110
          .         .         .         .         .         .
CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG
---------+---------+---------+---------+---------+---------+
GGCGTCGACTGTAGCAAGAATCATTTTGTCGTCTTTGTCTACACGAAAAATTATTTCACC
 A  A  D  I  V  L  S  K  T  A  E  T  D  V  L  F  N  K  V  E
         1130              1150              1170
          .         .         .         .         .         .
AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA
---------+---------+---------+---------+---------+---------+
TTAGTAGACCTTTCTTAATAAGTTTACAACTATGAAGTCAGTAAGGTAAGGTTTTACTAT
 S  S  G  K  N  Y  S  N  V  D  T  S  V  I  P  F  Q  N  D  M
```

FIG. 2C

```
        1190                1210                1230
         .         .         .         .         .         .
TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG
---------+---------+---------+---------+---------+---------+
ACGTATTACTACTTAGACCTTTTTTGTGACTACTAACAAATTTAGTGGTCTATTCATAAC
  H  N  D  E  S  G  K  N  T  D  D  C  L  N  H  Q  I  S  I  G
        1250                1270                1290
         .         .         .         .         .         .
GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA
---------+---------+---------+---------+---------+---------+
CACTGAAACCAATACCAGTAACATCATCACTTTAAAGATTGTAACTATTTTTGTGATTCT
  D  F  G  Y  G  H  C  S  S  E  I  S  N  I  D  K  N  T  K  N
        1310                1330                1350
         .         .         .         .         .         .
ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA
---------+---------+---------+---------+---------+---------+
TACGTAAAGTCCTGTAAAGTTACTCATTACATAGTACCCTCTTGAGAGTCTGCCTTATAT
  A  F  Q  D  I  S  M  S  N  V  S  W  E  N  S  Q  T  E  Y  S
        1370                1390                1410
         .         .         .         .         .         .
GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC
---------+---------+---------+---------+---------+---------+
CATTTTGAACAAAATATTCAAGGCAATTCGTGTGGGTCAGTCTTTTACCGTTATTTCTGG
  K  T  C  F  I  S  S  V  K  H  T  Q  S  E  N  G  N  K  D  H
        1430                1450                1470
         .         .         .         .         .         .
ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT
---------+---------+---------+---------+---------+---------+
TATATCTACTCTCACCCCTTTTACTCCTTCTTCGTCCAGAACTTTTGAGAAGCCTTTAAA
  I  D  E  S  G  E  N  E  E  E  A  G  L  E  N  S  S  E  I  S
        1490                1510                1530
         .         .         .         .         .         .
CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC
---------+---------+---------+---------+---------+---------+
GACGTCTACTCACCTCGTCCCCTTTATATGAATTTTTAAGTCACCCTCTCTTATAACTTG
  A  D  E  W  S  R  G  N  I  L  K  N  S  V  G  E  N  I  E  P
        1550                1570                1590
         .         .         .         .         .         .
CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC
---------+---------+---------+---------+---------+---------+
GACACTTTTAAAATCACGGACTTTTTTCAAATGGTACATTTCATTCATTATTATTAATAG
  V  K  I  L  V  P  E  K  S  L  P  C  K  V  S  N  N  N  Y  P
```

FIG. 2D

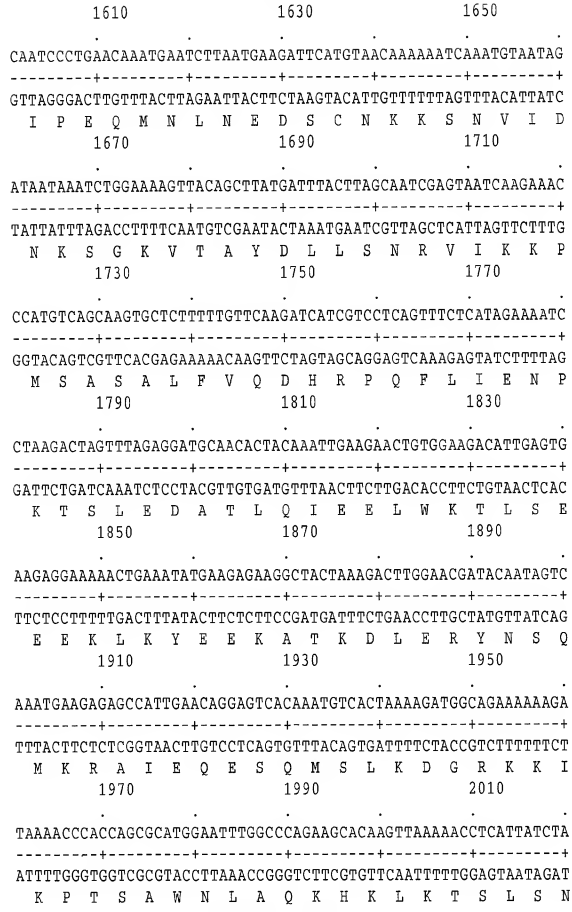
1610
1630
1650
CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG
GTTAGGGACTTGTTTACTTAGAATTACTTCTAAGTACATTGTTTTTTAGTTTACATTATC
I P E Q M N L N E D S C N K K S N V I D
1670
1690
1710
ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC
TATTATTTAGACCTTTTCAATGTCGAATACTAAATGAATCGTTAGCTCATTAGTTCTTTG
N K S G K V T A Y D L L S N R V I K K P
1730
1750
1770
CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC
GGTACAGTCGTTCACGAGAAAAACAAGTTCTAGTAGCAGGAGTCAAAGAGTATCTTTTAG
M S A S A L F V Q D H R P Q F L I E N P
1790
1810
1830
CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG
GATTCTGATCAAATCTCCTACGTTGTGATGTTTAACTTCTTGACACCTTCTGTAACTCAC
K T S L E D A T L Q I E E L W K T L S E
1850
1870
1890
AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC
TTCTCCTTTTTGACTTTATACTTCTCTTCCGATGATTTCTGAACCTTGCTATGTTATCAG
E E K L K Y E E K A T K D L E R Y N S Q
1910
1930
1950
AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAGA
TTTACTTCTCTCGGTAACTTGTCCTCAGTGTTTACAGTGATTTTCTACCGTCTTTTTTCT
M K R A I E Q E S Q M S L K D G R K K I
1970
1990
2010
TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA
ATTTTGGGTGGTCGCGTACCTTAAACCGGGTCTTCGTGTTCAATTTTTGGAGTAATAGAT
K P T S A W N L A Q K H K L K T S L S N

FIG. 2E

```
        2030                2050                2070
         .         .         .         .         .         .
ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA
---------+---------+---------+---------+---------+---------+
TAGTTGGTTTTGAACTACTTGAGGAAGTCAGGGTTTAACTTTTTTCTTCCTCAGTTTTAT
  Q  P  K  L  D  E  L  L  Q  S  Q  I  E  K  R  R  S  Q  N  I
        2090                2110                2130
         .         .         .         .         .         .
TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA
---------+---------+---------+---------+---------+---------+
AATTTTACCATGTCTAGGGGAAAAGATACTTTTTGAATTTTTATTTAAAATTCTTTGTTT
  K  M  V  Q  I  P  F  S  M  K  N  L  K  I  N  F  K  K  Q  N
        2150                2170                2190
         .         .         .         .         .         .
ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG
---------+---------+---------+---------+---------+---------+
TGTTTCAACTGAATCTTCTCTTCCTACTTGGAACGAACTAGGTGTTAGAGTCCAAAGGAC
  K  V  D  L  E  E  K  D  E  P  C  L  I  H  N  L  R  F  P  D
        2210                2230                2250
         .         .         .         .         .         .
ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG
---------+---------+---------+---------+---------+---------+
TACGTACCGATTACTGTAGGTTTTGTCTCCATTACAATAATTTAGGTATATCTCATCTTC
  A  W  L  M  T  S  K  T  E  V  M  L  L  N  P  Y  R  V  E  E
        2270                2290                2310
         .         .         .         .         .         .
AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA
---------+---------+---------+---------+---------+---------+
TTCGGGACGATAAATTTTCTGAAGAACTCTTAGTATTTGAAGGACGTCTCGGTGACCTTT
  A  L  L  F  K  R  L  L  E  N  H  K  L  P  A  E  P  L  E  K
        2330                2350                2370
         .         .         .         .         .         .
AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA
---------+---------+---------+---------+---------+---------+
TCGGTTAATACAATTGTCTCTCAGAAAAATTACCTAGAGTAATAAATCTGCAAAATATAT
  P  I  M  L  T  E  S  L  F  N  G  S  H  Y  L  D  V  L  Y  K
        2390                2410                2430
         .         .         .         .         .         .
AAATGACAGCAGATGACCAAAGATACAGTGGATCAACTTACCTGTCTGATCCTCGTCTTA
---------+---------+---------+---------+---------+---------+
TTTACTGTCGTCTACTGGTTTCTATGTCACCTAGTTGAATGGACAGACTAGGAGCAGAAT
  M  T  A  D  D  Q  R  Y  S  G  S  T  Y  L  S  D  P  R  L  T
```

FIG. 2F


2450
2470
2490
CAGCGAATGGTTTCAAGATAAAATTGATACCAGGAGTTTCAATTACTGAAAATTACTTGG
GTCGCTTACCAAAGTTCTATTTTAACTATGGTCCTCAAAGTTAATGACTTTTAATGAACC
A N G F K I K L I P G V S I T E N Y L E
2510
2530
2550
AAATAGAAGGAATGGCTAATTGTCTCCCATTCTATGGAGTAGCAGATTTAAAAGAAATTC
TTTATCTTCCTTACCGATTAACAGAGGGTAAGATACCTCATCGTCTAAATTTTCTTTAAG
I E G M A N C L P F Y G V A D L K E I L
2570
2590
2610
TTAATGCTATATTAAACAGAAATGCAAAGGAAGTTTATGAATGTAGACCTCGCAAAGTGA
AATTACGATATAATTTGTCTTTACGTTTCCTTCAAATACTTACATCTGGAGCGTTTCACT
N A I L N R N A K E V Y E C R P R K V I
2630
2650
2670
TAAGTTATTTAGAGGGAGAAGCAGTGCGTCTATCCAGACAATTACCCATGTACTTATCAA
ATTCAATAAATCTCCCTCTTCGTCACGCAGATAGGTCTGTTAATGGGTACATGAATAGTT
S Y L E G E A V R L S R Q L P M Y L S K
2690
2710
2730
AAGAGGACATCCAAGACATTATCTACAGAATGAAGCACCAGTTTGGAAATGAAATTAAAG
TTCTCCTGTAGGTTCTGTAATAGATGTCTTACTTCGTGGTCAAACCTTTACTTTAATTTC
E D I Q D I I Y R M K H Q F G N E I K E
2750
2770
2790
AGTGTGTTCATGGTCGCCCATTTTTTCATCATTTAACCTATCTTCCAGAAACTACATGAT
TCACACAAGTACCAGCGGGTAAAAAAGTAGTAAATTGGATAGAAGGTCTTTGATGTACTA
C V H G R P F F H H L T Y L P E T T *
2810
2830
2850
TAAATATGTTTAAGAAGATTAGTTACCATTGAAATTGGTTCTGTCATAAAACAGCATGAG
ATTTATACAAATTCTTCTAATCAATGGTAACTTTAACCAAGACAGTATTTTGTCGTACTC

FIG. 2G

```
      2870                2890                2910
         .         .         .         .         .         .
TCTGGTTTTAAATTATCTTTGTATTATGTGTCACATGGTTATTTTTTAAATGAGGATTCA
---------+---------+---------+---------+---------+---------+
AGACCAAAATTTAATAGAAACATAATACACAGTGTACCAATAAAAAATTTACTCCTAAGT
      2930               2950                 2970
         .         .         .         .         .         .
CTGACTTGTTTTTATATTGAAAAAAGTTCCACGTATTGTAGAAAACGTAAATAAACTAAT
---------+---------+---------+---------+---------+---------+
GACTGAACAAAAATATAACTTTTTTCAAGGTGCATAACATCTTTTGCATTTATTTGATTA

AAC
---
TTG
```

FIG. 2H

Polynucleotide and deduced amino acid sequence of hMLH3

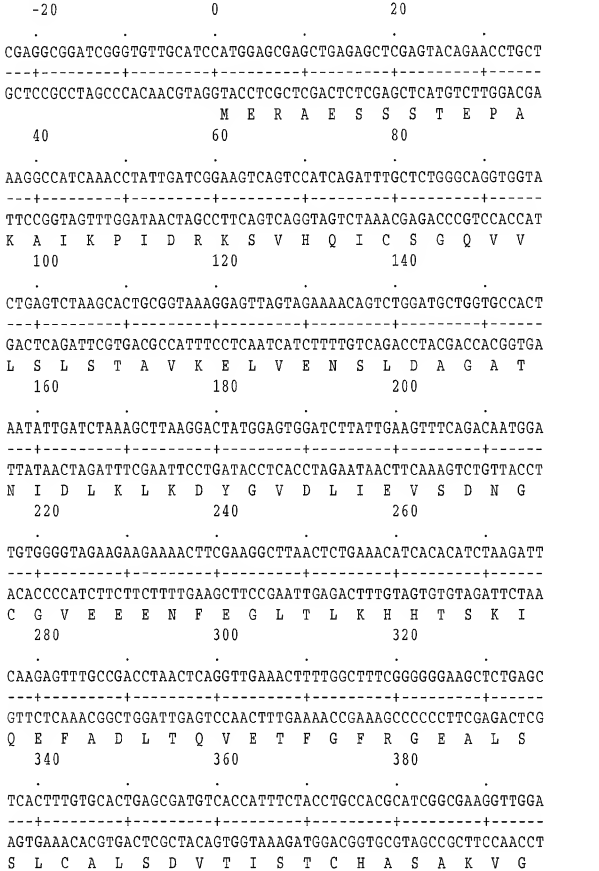

FIG. 3A

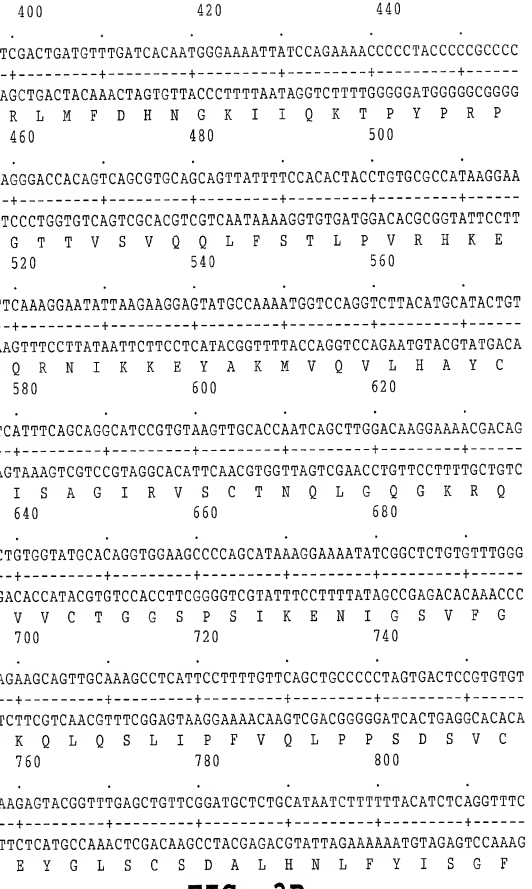

FIG. 3B

```
   820                 840                 860
   .         .         .         .         .         .
ATTTCACAATGCACGCATGGAGTTGGAAGGAGTTCAACAGACAGACAGTTTTTCTTTATC
---+---------+---------+---------+---------+---------+---------+------
TAAAGTGTTACGTGCGTACCTCAACCTTCCTCAAGTTGTCTGTCTGTCAAAAAGAAATAG
I  S  Q  C  T  H  G  V  G  R  S  S  T  D  R  Q  F  F  F  I
   880                 900                 920
   .         .         .         .         .         .
AACCGGCGGCCTTGTGACCCAGCAAAGGTCTGCAGACTCGTGAATGAGGTCTACCACATG
---+---------+---------+---------+---------+---------+---------+------
TTGGCCGCCGGAACACTGGGTCGTTTCCAGACGTCTGAGCACTTACTCCAGATGGTGTAC
N  R  R  P  C  D  P  A  K  V  C  R  L  V  N  E  V  Y  H  M
   940                 960                 980
   .         .         .         .         .         .
TATAATCGACACCAGTATCCATTTGTTGTTCTTAACATTTCTGTTGATTCAGAATGCGTT
---+---------+---------+---------+---------+---------+---------+------
ATATTAGCTGTGGTCATAGGTAAACAACAAGAATTGTAAAGACAACTAAGTCTTACGCAA
Y  N  R  H  Q  Y  P  F  V  V  L  N  I  S  V  D  S  E  C  V
   1000                1020                1040
   .         .         .         .         .         .
GATATCAATGTTACTCCAGATAAAAGGCAAATTTTGCTACAAGAGGAAAAGCTTTTGTTG
---+---------+---------+---------+---------+---------+---------+------
CTATAGTTACAATGAGGTCTATTTTCCGTTTAAAACGATGTTCTCCTTTTCGAAAACAAC
D  I  N  V  T  P  D  K  R  Q  I  L  L  Q  E  E  K  L  L  L
   1060                1080                1100
   .         .         .         .         .         .
GCAGTTTTAAAGACCTCTTTGATAGGAATGTTTGATAGTGATGTCAACAAGCTAAATGTC
---+---------+---------+---------+---------+---------+---------+------
CGTCAAAATTTCTGGAGAAACTATCCTTACAAACTATCACTACAGTTGTTCGATTTACAG
A  V  L  K  T  S  L  I  G  M  F  D  S  D  V  N  K  L  N  V
   1120                1140                1160
   .         .         .         .         .         .
AGTCAGCAGCCACTGCTGGATGTTGAAGGTAACTTAATAAAAATGCATGCAGCGGATTTG
---+---------+---------+---------+---------+---------+---------+------
TCAGTCGTCGGTGACGACCTACAACTTCCATTGAATTATTTTTACGTACGTCGCCTAAAC
S  Q  Q  P  L  L  D  V  E  G  N  L  I  K  M  H  A  A  D  L
   1180                1200                1220
   .         .         .         .         .         .
GAAAAGCCCATGGTAGAAAAGCAGGATCAATCCCCTTCATTAAGGACTGGAGAAGAAAAA
---+---------+---------+---------+---------+---------+---------+------
CTTTTCGGGTACCATCTTTTCGTCCTAGTTAGGGGAAGTAATTCCTGACCTCTTCTTTTT
E  K  P  M  V  E  K  Q  D  Q  S  P  S  L  R  T  G  E  E  K
```

FIG. 3C

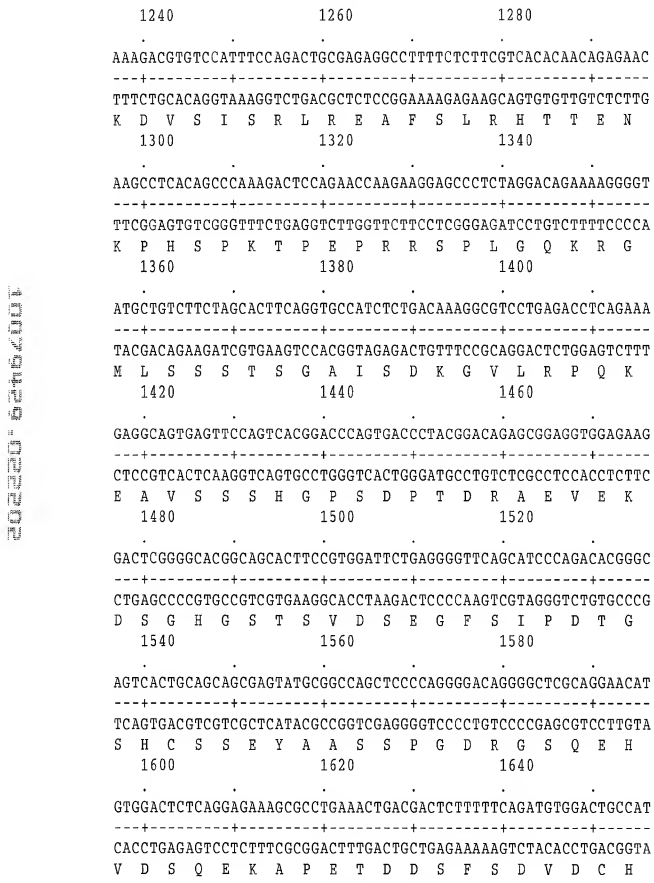
1240
1260
1280
AAAGACGTGTCCATTTCCAGACTGCGAGAGGCCTTTTCTCTTCGTCACACAACAGAGAAC
TTTCTGCACAGGTAAAGGTCTGACGCTCTCCGGAAAAGAGAAGCAGTGTGTTGTCTCTTG
K D V S I S R L R E A F S L R H T T E N
1300
1320
1340
AAGCCTCACAGCCCAAAGACTCCAGAACCAAGAAGGAGCCCTCTAGGACAGAAAAGGGGT
TTCGGAGTGTCGGGTTTCTGAGGTCTTGGTTCTTCCTCGGGAGATCCTGTCTTTTCCCCA
K P H S P K T P E P R R S P L G Q K R G
1360
1380
1400
ATGCTGTCTTCTAGCACTTCAGGTGCCATCTCTGACAAAGGCGTCCTGAGACCTCAGAAA
TACGACAGAAGATCGTGAAGTCCACGGTAGAGACTGTTTCCGCAGGACTCTGGAGTCTTT
M L S S S T S G A I S D K G V L R P Q K
1420
1440
1460
GAGGCAGTGAGTTCCAGTCACGGACCCAGTGACCCTACGGACAGAGCGGAGGTGGAGAAG
CTCCGTCACTCAAGGTCAGTGCCTGGGTCACTGGGATGCCTGTCTCGCCTCCACCTCTTC
E A V S S S H G P S D P T D R A E V E K
1480
1500
1520
GACTCGGGGCACGGCAGCACTTCCGTGGATTCTGAGGGGTTCAGCATCCCAGACACGGGC
CTGAGCCCCGTGCCGTCGTGAAGGCACCTAAGACTCCCCAAGTCGTAGGGTCTGTGCCCG
D S G H G S T S V D S E G F S I P D T G
1540
1560
1580
AGTCACTGCAGCAGCGAGTATGCGGCCAGCTCCCCAGGGGACAGGGGCTCGCAGGAACAT
TCAGTGACGTCGTCGCTCATACGCCGGTCGAGGGGTCCCCTGTCCCCGAGCGTCCTTGTA
S H C S S E Y A A S S P G D R G S Q E H
1600
1620
1640
GTGGACTCTCAGGAGAAAGCGCCTGAAACTGACGACTCTTTTTCAGATGTGGACTGCCAT
CACCTGAGAGTCCTCTTTCGCGGACTTTGACTGCTGAGAAAAAGTCTACACCTGACGGTA
V D S Q E K A P E T D D S F S D V D C H

FIG. 3D

```
           1660                1680                1700
             .         .         .         .         .         .
   TCAAACCAGGAAGATACCGGATGTAAATTTCGAGTTTTGCCTCAGCCAACTAATCTCGCA
   ---+---------+---------+---------+---------+---------+------
   AGTTTGGTCCTTCTATGGCCTACATTTAAAGCTCAAAACGGAGTCGGTTGATTAGAGCGT
   S  N  Q  E  D  T  G  C  K  F  R  V  L  P  Q  P  T  N  L  A
          1720                1740                1760
             .         .         .         .         .         .
   ACCCCAAACACAAAGCGTTTTAAAAAAGAAGAAATTCTTTCCAGTTCTGACATTTGTCAA
   ---+---------+---------+---------+---------+---------+------
   TGGGGTTTGTGTTTCGCAAAATTTTTTCTTCTTTAAGAAAGGTCAAGACTGTAAACAGTT
   T  P  N  T  K  R  F  K  K  E  E  I  L  S  S  S  D  I  C  Q
          1780                1800                1820
             .         .         .         .         .         .
   AAGTTAGTAAATACTCAGGACATGTCAGCCTCTCAGGTTGATGTAGCTGTGAAAATTAAT
   ---+---------+---------+---------+---------+---------+------
   TTCAATCATTTATGAGTCCTGTACAGTCGGAGAGTCCAACTACATCGACACTTTTAATTA
   K  L  V  N  T  Q  D  M  S  A  S  Q  V  D  V  A  V  K  I  N
          1840                1860                1880
             .         .         .         .         .         .
   AAGAAAGTTGTGCCCCTGGACTTTTCTATGAGTTCTTTAGCTAAACGAATAAAGCAGTTA
   ---+---------+---------+---------+---------+---------+------
   TTCTTTCAACACGGGGACCTGAAAAGATACTCAAGAAATCGATTTGCTTATTTCGTCAAT
   K  K  V  V  P  L  D  F  S  M  S  S  L  A  K  R  I  K  Q  L
          1900                1920                1940
             .         .         .         .         .         .
   CATCATGAAGCACAGCAAAGTGAAGGGGAACAGAATTACAGGAAGTTTAGGGCAAAGATT
   ---+---------+---------+---------+---------+---------+------
   GTAGTACTTCGTGTCGTTTCACTTCCCCTTGTCTTAATGTCCTTCAAATCCCGTTTCTAA
   H  H  E  A  Q  Q  S  E  G  E  Q  N  Y  R  K  F  R  A  K  I
          1960                1980                2000
             .         .         .         .         .         .
   TGTCCTGGAGAAAATCAAGCAGCCGAAGATGAACTAAGAAAAGAGATAAGTAAAACGATG
   ---+---------+---------+---------+---------+---------+------
   ACAGGACCTCTTTTAGTTCGTCGGCTTCTACTTGATTCTTTTCTCTATTCATTTTGCTAC
   C  P  G  E  N  Q  A  A  E  D  E  L  R  K  E  I  S  K  T  M
          2020                2040                2060
             .         .         .         .         .         .
   TTTGCAGAAATGGAAATCATTGGTCAGTTTAACCTGGGATTTATAATAACCAAACTGAAT
   ---+---------+---------+---------+---------+---------+------
   AAACGTCTTTACCTTTAGTAACCAGTCAAATTGGACCCTAAATATTATTGGTTTGACTTA
   F  A  E  M  E  I  I  G  Q  F  N  L  G  F  I  I  T  K  L  N
```

FIG. 3E

```
   2080                2100                2120
    .         .         .         .         .         .
GAGGATATCTTCATAGTGGACCAGCATGCCACGGACGAGAAGTATAACTTCGAGATGCTG
---+---------+---------+---------+---------+---------+---------+------
CTCCTATAGAAGTATCACCTGGTCGTACGGTGCCTGCTCTTCATATTGAAGCTCTACGAC
E  D  I  F  I  V  D  Q  H  A  T  D  E  K  Y  N  F  E  M  L
   2140                2160                2180
    .         .         .         .         .         .
CAGCAGCACACCGTGCTCCAGGGGCAGAGGCTCATAGCACCTCAGACTCTCAACTTAACT
---+---------+---------+---------+---------+---------+---------+------
GTCGTCGTGTGGCACGAGGTCCCCGTCTCCGAGTATCGTGGAGTCTGAGAGTTGAATTGA
Q  Q  H  T  V  L  Q  G  Q  R  L  I  A  P  Q  T  L  N  L  T
   2200                2220                2240
    .         .         .         .         .         .
GCTGTTAATGAAGCTGTTCTGATAGAAAATCTGGAAATATTTAGAAAGAATGGCTTTGAT
---+---------+---------+---------+---------+---------+---------+------
CGACAATTACTTCGACAAGACTATCTTTTAGACCTTTATAAATCTTTCTTACCGAAACTA
A  V  N  E  A  V  L  I  E  N  L  E  I  F  R  K  N  G  F  D
   2260                2280                2300
    .         .         .         .         .         .
TTTGTTATCGATGAAAATGCTCCAGTCACTGAAAGGGCTAAACTGATTTCCTTGCCAACT
---+---------+---------+---------+---------+---------+---------+------
AAACAATAGCTACTTTTACGAGGTCAGTGACTTTCCCGATTTGACTAAAGGAACGGTTGA
F  V  I  D  E  N  A  P  V  T  E  R  A  K  L  I  S  L  P  T
   2320                2340                2360
    .         .         .         .         .         .
AGTAAAAACTGGACCTTCGGACCCCAGGACGTCGATGAACTGATCTTCATGCTGAGCGAC
---+---------+---------+---------+---------+---------+---------+------
TCATTTTTGACCTGGAAGCCTGGGGTCCTGCAGCTACTTGACTAGAAGTACGACTCGCTG
S  K  N  W  T  F  G  P  Q  D  V  D  E  L  I  F  M  L  S  D
   2380                2400                2420
    .         .         .         .         .         .
AGCCCTGGGGTCATGTGCCGGCCTTCCCGAGTCAAGCAGATGTTTGCCTCCAGAGCCTGC
---+---------+---------+---------+---------+---------+---------+------
TCGGGACCCCAGTACACGGCCGGAAGGGCTCAGTTCGTCTACAAACGGAGGTCTCGGACG
S  P  G  V  M  C  R  P  S  R  V  K  Q  M  F  A  S  R  A  C
    2440                2460                2480
    .         .         .         .         .         .
CGGAAGTCGGTGATGATTGGGACTGCTCTTAACACAAGCGAGATGAAGAAACTGATCACC
---+---------+---------+---------+---------+---------+---------+------
GCCTTCAGCCACTACTAACCCTGACGAGAATTGTGTTCGCTCTACTTCTTTGACTAGTGG
R  K  S  V  M  I  G  T  A  L  N  T  S  E  M  K  K  L  I  T
```

FIG. 3F

```
   2500                2520                2540
    .         .         .         .         .         .
CACATGGGGGAGATGGACCACCCCTGGAACTGTCCCCATGGAAGGCCAACCATGAGACAC
---+---------+---------+---------+---------+---------+------
GTGTACCCCCTCTACCTGGTGGGGACCTTGACAGGGGTACCTTCCGGTTGGTACTCTGTG
H  M  G  E  M  D  H  P  W  N  C  P  H  G  R  P  T  M  R  H
   2560                2580                2600
    .         .         .         .         .         .
ATCGCCAACCTGGGTGTCATTTCTCAGAACTGACCGTAGTCACTGTATGGAATAATTGGT
---+---------+---------+---------+---------+---------+------
TAGCGGTTGGACCCACAGTAAAGAGTCTTGACTGGCATCAGTGACATACCTTATTAACCA
I  A  N  L  G  V  I  S  Q  N  *
   2620                2640                2660
    .         .         .         .         .         .
TTTATCGCAGATTTTTATGTTTTGAAAGACAGAGTCTTCACTAACCTTTTTTGTTTTAAA
---+---------+---------+---------+---------+---------+------
AAATAGCGTCTAAAAATACAAAACTTTCTGTCTCAGAAGTGATTGGAAAAAACAAAATTT
   2680                2700                2720
    .         .         .         .         .         .
ATGAAACCTGCTACTTAAAAAAAATACACATCACACCCATTTAAAAGTGATCTTGAGAAC
---+---------+---------+---------+---------+---------+------
TACTTTGGACGATGAATTTTTTTTATGTGTAGTGTGGGTAAATTTTCACTAGAACTCTTG
   2740
    .
CTTTTCAAACC
---+-------
GAAAAGTTTGG
```

FIG. 3G

```
YPMS1   mfhhienllietekrckqkeqryipvkylfsmtqIHQINDIDVHRITSGQVITDLTTAVKELVDNSIDANANQIEIIFKD
HMLH2   ----------------------------------MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLEN
HMLH3   meraessstepaka--------------------IKPIDRKSVHQICSGQVVLSLSTAVKELVENSLDAGATNIDLKLKD

YPMS1   YGLESIECSDNGDGIDPSNYEFLALKHYTSKIAKFQDVAKVQTLGFRGEALSSLCGIAKLSVITTTSPPK-ADKLEYDMV
HMLH2   YGFDKIEVRDNGEGIKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQYVLDGS
HMLH3   YGVDLIEVSDNGCGVEEENFEGLTLKHHTSKIQEFADLTQVETFGFRGEALSSLCALSDVTISTCHASAKVGTRLMFDHN

YPMS1   GHITSKTTTSRNKGTTVLVSQLFHNLPVRQKEFSKTfkrqftkcltviqgyaiinaaikfsvwnitpkgkknlilstmrn
HMLH2   GHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAkkckdeikkiqdllmsfgilkpdlrivfvhnkaviwqksrvsdh
HMLH3   GKIIQKTPYPRPRGTTVSVQQLFSTLPVRHKEFQRNikkeyakmvqvlhayciisagirvsctnqlgqgkrqpvvctggs

YPMS1   ssmrknissvfgaggmrgleevdlvldlnpfknrmlgkytddpdfidldykirvkgyisqnsfgcgrNSKDRQFIYVNKR
HMLH2   kmalmsvlgtavmnnmesfqyhseesqiylsgflpkcdadhsftsl----------STPERSFIFINSR
HMLH3   psikenigsvfgqkqlqslipfvqlppsdsvceeyglscsdalhnlfyisgfisqcthgvgr-----SSTDRQFFFINRR
```

FIG. 4A

```
YPMS1   PVEYSTLLKCCNEVVKTFNNVQ----FPAVFLNLELPMSLIDVNVTPDKRVILLHNERAVIDIFKTTLSDYYNrqelalp
HMLH2   PVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKIDVPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGplpstns
HMLH3   RCDPAKVCRLVNEVYHMYNRHQ----YPFVVLNISVDSECVDINVTPDKRQILLQEEKLLLAVLKTSLIGMFDsdvnkln

YPMS1   krmcsqseqqaqkrlktevfddrstthesdnenyhtarsesnqsnhahfnsttgvidksngteltsvmdgnytnvtdvig
HMLH2   yennktdvsaadivlsktaetdvlfnkvessgknysnvdtsvipfqndmhndesgkntddclnhqisigdfgyghcssei
HMLH3   vsqqplldvegnlikmhaadlekpmvekqdqspslrtgeekkdvsisrlreafslrhttenkphspktpeprrsplgkkr

YPMS1   secevsvdssvvldegnsstptkklpsiktdsqnlsdlnlnnfsnpefqnitspdkarslekvveepvyfdidgekfqek
HMLH2   snidkntknafqdismsnvswensqteysktcfissvkhtqsengnkdhidesgeneeeaglensseisadewsrgnilk
HMLH3   gmlssstsgaisdkgvlrpqkeavsssshgpsdptdraevekdsghgstsvdsegfsipdtgshcsseyaasspgdrgsqe

YPMS1   avlsqadglvfvdnechehtndcchqerrgstdteqddeadsiyaeiepveinvrtplknsrksiskdnyrslsdglthr
HMLH2   nsvgeniepvkilvpekslpckvsnnnypipeqmnlnedscnkksnvidnksgkvtaydllsnrvikkpmsasalfvqdh
HMLH3   hvdsqekapetddsfsdvdchsnqedtgckfrvlpqptnlatpntkrfkkeeilssdicqklvntqdmsasqvdvavki
```

FIG. 4B

```
YPMS1   kfedeileynlstknfkeisknqkqmssliskrkseaqeniiknkdeledfeqgekyltltvskndfkkmevvgqfnlgf
HMLH2   rpqflienpktsledatlqieelwktlseeeklkyeekatkdlerynsqmkraieqesqmslkdgrkkikptsawnlaqk
HMLH3   nkkvvpldfsmsslakrikqlhheaqqsegeqnyrkfrakicpgenqaaedelrkeisktmfaemeiigqfnlgfiitkl

YPMS1   iivtrkvdnksdlfivdqhasdekynfetlqavtvfksqkliipqpvelsvidelvvldnlpvfekngfklkideeeefg
HMLH2   hklktslsnqpkldellqsqiekrrsqnikmvqipfsmknlkinfkkqnkvdleekdepclihnlrfpdawlmtsktevm
HMLH3   nedifivdqhatdekynfemlqqhtvlqgqrliapqtlnltavneavlienleifrkngfdfvidenapvteraklislp

YPMS1   srvkllslptskqtlfdlgdfnelihlikedgglrrdni----------------------------------------
HMLH2   llnpyrveeallfkrllenhklpaeplekpimlteslfngshyldvlykmtaddqrysgstylsdprltangfkiklipg
HMLH3   tsknwtfgpqdvdelifmlsdspgvmc----------------------------------------------------

YPMS1   ---------------------------------------- RCSKIRSMFAMRACRSSIMIGKPLNKKTMTRVVHNLs
HMLH2   vsitenyleiegmanclpfygvadlkeilnailnrnakevyecRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDILYRMk
HMLH3   ---------------------------------------- RPSRVKQMFASRACRKSVMIGTALNTSEMKKLITHMg

YPMS1   eldkpw--NCPHGRPTMRHLMELrdwssfskdyei
HMLH2   hqfgneikECVHGRPFFHHLTYLpett--------
HMLH3   emdhpw--NCPHGRPTMRHIANLgvisqn------
```

FIG. 4C

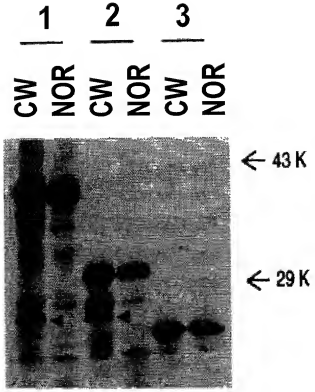

FIG. 5A

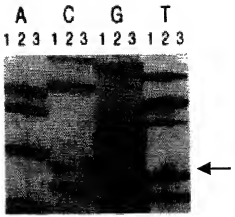

FIG. 5B

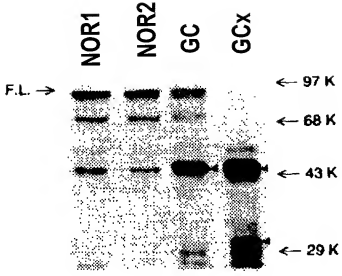
NOR1
NOR2
GC
GCx
F.L. →
← 97 K
← 68 K
← 43 K
← 29 K

FIG. 6A

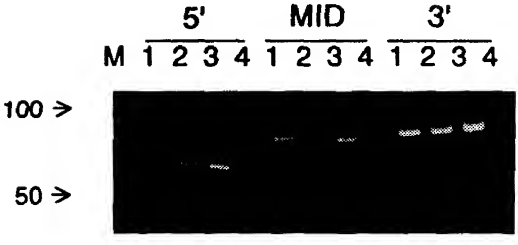
5'
MID
3'
M 1 2 3 4 1 2 3 4 1 2 3 4
100 →
50 →

FIG. 6B